DOCUMENT RESUME

ED 079 175 SO 005 831

AUTHOR Ehman, Lee H.
TITLE Teaching Behavior Patterns and Personal Teacher Attributes Related to Success and Failure in Specific Social Studies Teaching Tasks.
PUB DATE 1 Mar 73
NOTE 17p.; Paper presented at annual Meeting of American Educational Research Association (New Orleans, Louisiana, February 25-March 1, 1973)

EDRS PRICE MF-$0.65 HC-$3.29
DESCRIPTORS Bibliographic Citations; *Educational Research; Microteaching; Models; Research Projects; Secondary Grades; *Social Studies; Speeches; Student Reaction; *Student Teacher Relationship; *Teacher Behavior; *Teacher Characteristics; Teaching Techniques
IDENTIFIERS *American Educational Research Association

ABSTRACT

This paper examines the relationships between both low- and high-inference teaching process variables and success in specific teaching tasks performed in a laboratory setting. Two minor objectives are also included: teacher characteristics are related to success in the teaching tasks, and the relationships between low- and high-inference teaching process variables are analyzed. The intermediate effects, overt pupil responses and covert pupil responses, are used as criteria of teaching success in this process-product study. The research method is detailed and the findings of the study are summarized in tabular form. (Author/SHM)

ED 079175

# TEACHING BEHAVIOR PATTERNS AND PERSONAL TEACHER ATTRIBUTES RELATED TO SUCCESS AND FAILURE IN SPECIFIC SOCIAL STUDIES TEACHING TASKS

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

Lee H. Ehman

Indiana University

SP005831

Presented at the annual meeting of the American Educational Research Association, New Orleans, February 25 - March 1, 1973.

## Objectives

This paper examines the relationships between both low-and high-inference teaching process variables and success in specific teaching tasks performed in a laboratory teaching setting. Two minor objectives are also included: teacher characteristics are related to success in the teaching tasks, and the relationships between low- and high-inference teaching process variables are analyzed.

Rosenshine and Furst, in their extensive review of process-product teaching research, distinguished between low inference variables, or those generally derived from observation instruments and which focus on specific teaching behavior, and high inference variables, generally obtained from rating systems and which lack the specificity of low-inference measures.[1] Both types of variables have been used in studies of teacher performance, although Rosenshine and Furst conclude that high-inference variables tend to be more clearly related to student achievement than low-inference variables. This study will use both types of variables in combination to assess their relative contribution to the success of specific teaching tasks.

Rosenshine and Furst also pointed out that little research has focussed on the low-inference components which comprise high-inference ratings.[2] For example, what specific teacher behaviors are associated with the "clarity" of the teaching task as presented by the teacher to students? This question will also be addressed in the paper.

Perhaps the oldest (and least fruitful) "branch" of research on teacher behavior is that which has attempted to link teacher characteristics to success in teaching. This paper will include three such "presage" characteristic -- social science intellectual skill development, cognitive complexity, and attitudes toward teacher control -- in the analysis of teaching task success.

This paper does not focus on student intellectual achievement as the criterion of teaching success. Rather, "intermediate effects," as discussed by Biddle, are used as criteria.[3] Biddle conceptualized classroom process situations as having three sets of teacher-and student-related variables, as seen in Figure 1. (This

Figure 1

| Classroom Situations<br>a. Physical Environment<br>b. Social Incidents | | | |
|---|---|---|---|
| Teacher Properties | Teacher Behaviors | Intermediate Effects | Long-Term Consequences |
| a. Skills<br>b. Motives<br>c. Habits<br>d. Knowledge | a. Traits<br>b. Responses to Environment | a. Overt Pupil Responses<br>b. Covert Pupil Responses | a. Achievement or Adjustment of Pupils<br>b. New Ideas in Education<br>c. Aggrandizement of the Profession |

figure is only a part of Biddle's seven-variable model.) Biddle argues that there are intermediate effects, in terms of pupil responses in classrooms, that may mediate between the teaching act and the longer-term learning outcomes. These intermediate effects are often left out of process-product studies, but are studied here. The other two teacher variable sets in Biddle's model are also accounted for in this study, with the "teacher properties" of cognitive complexity, social science intellectual skill, and attitudes toward control, and both high- and low-inference teacher behaviors.

The intermediate effects studied here are closely related to the teaching tasks being performed. These tasks include teaching concepts, eliciting student hypotheses concerning social science problems, and analyzing valuational issues. These tasks are part of an overall social studies inquiry teaching strategy. The three elements do not exhaust the possible facets of inquiry teaching, but they do represent major parts of prominent inquiry teaching theories.

Procedure

Data used in this research were collected in a pre-service secondary social studies teacher training laboratory. Connected with several sections of an undergraduate secondary methods course, this laboratory enabled researchers to record large numbers of ten to fifteen minute lessons taught by the methods class pre-service teachers to groups of four to six ninth grade students who were hired for after-school work in the laboratory. Instructors in the methods classes assigned each student seven micro-teaching lessons as an integral part of the course. The first, taught only to peers, was aimed at warming up and familiarizing the methods student to the laboratory setting and mechanics of micro-teaching. These lessons were not used in the present study. The remaining six lessons were taught in pairs at about three-week intervals through the semester. In the first pair of lessons, the student invented a lesson in which one or more historical/social science concepts were taught. The content was selected by each student, but the teaching task was fixed. A plan was submitted, critiqued individually, and often revised; all of this took place before the teaching. Then the student taught the lesson to a group of ninth graders. He was critiqued by the instructor, who also played back for analysis selected parts of the video-taped lesson. Some students modified the lesson's structure or sequence according to the critique or his own ideas; others did not modify them. The content and task remained the same. Two days later, the same lesson was retaught under similar conditions to a new group of ninth grade students. Three weeks later, the next pair of lessons was taught, involving the task of eliciting student hypotheses in response to a problem chosen by the student. The third pair of lessons involved the analysis of a value proposition or controversial issue. Thus, there were three pairs of similar lessons taught under controlled conditions, in which all teachers had the same basic objectives assigned to them. In all, eighty-four students taught lessons which were recorded for analysis; in some cases, a full set of recorded lessons were not obtained because of mechanical or other difficulties. For the analysis which follows, all students for which data is available will be included, so that this number may vary somewhat across analyses.

Each lesson, with the exceptions noted, was recorded on an audio tape cassette recorder, and then coded according to a modified version of the Flanders verbal interaction analysis system.[4] The basic ten categories were used, and teacher's question (category 4,) student responses (category 8,) student-initiated statements (category 9,) and silence/confusion (category 10,) were modified. Categories 4, 8,

and 9, were subscripted according to the Aschner-Gallagher coding scheme.[5] Cognitive levels or modes implied by the verbal statements were coded into one of four categories: cognitive-memory; convergent; divergent; and evaluative. Figure 2 lists the twenty categories in the system.

Figure 2

MODIFIED FLANDERS VERBAL INTERACTION ANALYSIS SYSTEM

| Category Symbol | Category Label |
|---|---|
| 1 | Accepts feelings |
| 2 | Praises/encourages |
| 3 | Accepts/uses student ideas |
| 41 | Asks cognitive-memory question |
| 42 | Asks convergent question |
| 43 | Asks divergent question |
| 44 | Asks evaluative question |
| 5 | Lectures |
| 6 | Gives Directions |
| 7 | Criticizes |
| 81 | Student response-cognitive memory |
| 82 | " " convergent |
| 83 | " " divergent |
| 84 | " " evaluative |
| 91 | Student initiation-cognitive memory |
| 92 | " " convergent |
| 93 | " " divergent |
| 94 | " " evaluative |
| 01 | Productive silence or confusion |
| 02 | Unproductive silence or confusion |

Lessons were coded by designating the category of the predominant verbal behavior for each three-second period. Three persons coded a random sample of tapes in a random sequence, so that all lessons of the same type were not coded in blocks. Inter- and intra-coder reliabilities were estimated by having each coder work with criterion tapes at intervals throughout the project. Scott's coefficients for these reliability tests are presented in Table 1.

Table 1

SCOTT'S COEFFICIENTS FOR INTER- AND INTRA- CODER RELIABILITY

| | Inter-coder* | | | Intra-coder** | | |
|---|---|---|---|---|---|---|
| | 1-2 | 1-3 | 2-3 | 1 | 2 | 3 |
| Basic 10 categories | .79 | .82 | .82 | .78 | .79 | .78 |
| Categories 4, 5, 8, and 9 with subscripts (13 categories) | .75 | .79 | .80 | .80 | .79 | .78 |
| All categories (20 categories) | .63 | .68 | .72 | .71 | .71 | .73 |

* Means of 6 tests
** Means of 2 tests

In addition to these low-inference measurers, ratings of each tape were made by a rater who was not involved in the coding described above. This person listened to each audio tape in a different random order than that used by the coders, and rated the lessons on several dimensions, some of which were suggested by the Rosenshine and Furst review. The rater also made two or three simple frequency counts of teacher and student behaviors not included in the coding. Lessons were rated using a stanine scale on the following characteristic:

1. Clarity of the task students were expected to work on

2. Effectiveness of the lessons' introduction in capturing student interest

3. Teacher's enthusiasm

4. Teacher's businesslike (or task-oriented) manner

5. Clarity of the concept or intellectual problem being discussed

6. Overall student response quality in the lesson (did they seem to grasp the concept, were they able to hypothesize solutions or answers to the problems, or were they able to analyze the valuational problems?)

The following frequency counts were also made:

1. Number of teacher structuring statements

2. Number of different concept examples students articulated (for the concept lessons only)

3. Number of student hypotheses articulated (for the hypothesis lesson only)

Because there was only one rater, the reliability of the rating variables is a major problem in the study. The rater did rate 38 of the 445 lessons twice, and intra-coder reliability was estimated, using product-moment correlations on all measures. These correlations ranged from .59 to .96, with a mean of .75.

The overall student response quality rating for each lesson, plus the latter two frequency counts, will constitute the "intermediate effects," or critical outcomes for the research. The intra-coder reliabilities for these three variables were:

1. Overall student response quality .86
2. Number of different examples of concepts .96
3. Number of hypotheses generated by students .81

Finally, three paper and pencil instruments were administered to the preservice teachers. The first was the Harvey, Hurt, and Schroder Conceptual Systems Test, originally used with elementary school teachers.[6] The responses of the eighty-four teachers were factor-analyzed, and factor scores based on varimax factor score weights on five of the factors were generated for each teacher. The sixth factor was judged to be uninterpretable in the conceptual framework presented by the authors of the instrument. The five factors are:

1. Need for People
2. Belief in Divine Fate Control
3. Need to Help People
4. Need for Structure and Order
5. Anomie

Except for the one uninterpretable factor, the varimax factor structure was very similar to that reported in the original work done by Harvey, Hunt, and Schroder. Because the instrument purportedly tops dimensions of cognitive complexity, from concrete to abstract, it suggested a teacher characteristic which might be linked to success in performing inquiry teaching tasks. The second teacher trait measure used was the social science skills instrument, developed by Bowman.[7] Forty-two items test the teachers' abilities in the following areas:

1. Discrimination between hypotheses implied by data and those not implied by data.
2. Discrimination between questions pertinent to the validation of hypotheses and those not pertinent to this validation.
3. Judge whether analogies support or challenge positions on controversial issues.
4. Judge whether legal principles are violated or supported by potential policy decisions.
5. Discriminate between conclusions which follow from premises and those that do not follow.

These skills were judged relevant to the teaching tasks being performed. A total test score was computed for each teacher. The reliability of the total test was estimated at .65, using the Kuder-Richardson formula 21, and .69 for Cronbach's alpha. These reliabilities suggest weakness in the test.

Finally, an instrument assessing attitudes toward teacher control in classrooms, called the Teacher Control Altitude Scale, was also administered.[8] This thirty-item instrument was developed through judges' ratings of items discriminating between high and low controlling behavior by teachers in the classroom. A single total score was computed for each teacher. The KR-21 reliability estimate for this instrument is .78.

## Analysis

The analysis will consist of three sections: 1) a description of the three lesson types in terms of low-inference observation variables; 2) an analysis of the relationships between teaching task success and the teacher trait and behavior variables; 3) an examination of the low-inference components of high-inference teaching variables.

Lesson description. Different verbal processes were expected for the three lessons because the specific teaching tasks in each lesson varied considerably. The concept teaching lesson was expected to be more direct and teacher-oriented than the other two lesson types. The task of teaching a concept to a group of students who do not initially understand its meaning is relatively teacher-centered throughout. In contrast, the hypothesis-eliciting and value analysis tasks were expected to be teacher-focussed while the problem or issue is being introduced and framed, but there the focus should shift to the students, who are expected to generate hypotheses or the analysis and grounding value positions. This difference was expected to result in differences in the proportion of time in teacher-talk. The direction of the differences in teacher-talk proportions do bear out this expectation, but the magnitude is slight. The average proportions are .69, .67, and .64 for the concept, hypothesis, and values lessons, respectively. All three proportions are close to the average of 68 per cent observed by Flanders.[9] Bellack, et al, found an overall teacher-talk median of 72.6 per cent of lines in transcripts for sixty social studies lessons, and this also corresponds closely to the present findings.[10]

Differences in proportions within the several teacher-talk categories were also anticipated, but were not observed. Table 2 shows the startlingly similar proportions in the six teacher-talk categories in which behavior was observed. More lecturing and less use of student ideas was expected for the concept lesson than for the other two types, but the differences were trivial.

Table 2

MEAN PROPORTIONS* IN BASIC TEACHER-TALK CATEGORIES IN THREE LESSON TYPES

| Basic Category | Lesson Type: Concept | Hypothesis | Values |
|---|---|---|---|
| 1 Accepts feelings | (No behavior observed in this category) | | |
| 2 Praises/encourages | 5.94 | 5.35 | 3.35 |
| 3 Accepts/uses student ideas | 5.25 | 5.65 | 5.09 |
| 4 Asks questions | 13.10 | 13.66 | 14.39 |
| 5 Lectures | 35.45 | 33.49 | 33.84 |
| 6 Gives directions | 1.00 | 1.09 | 0.57 |
| 7 Criticizes | 0.12 | 0.15 | 0.03 |

* The proportions represent the time in the category divided by the total time in the lesson.

A difference in the balance between praise (category 2) and use of ideas (category 3) was also expected across the lesson types. In the concept-teaching lesson, where teacher cuing behavior was expected to be more prominent than in the other two lessons, simple praise was expected to be relatively greater than was of student ideas. Yet the balance between these categories is constant in all three lesson types.

No instances of accepting feelings were observed, and this may be due to the very short amount of interpersonal contact between teacher and students allowed by the micro-teaching laboratory setting. Of course, Flanders and others have reported a very low incidence of category 1 -- less than one per cent -- across the variety of subjects and situations. The proportion in category 7, Teacher Gives Criticism, was expected to be small because of the controlled laboratory situation, and this proved to be the case. Because the lessons were short and self-contained, rather few directions to students were called for, and this is supported by the very small proportions in category 6. In each lesson type, about half of the teacher talk, or one-third of the total lesson time, was devoted to teacher lecturing behavior, category 5. Explaining defining characteristics and giving examples of a concept seem to take about as much teacher talk as the problem-introduction and explaining process necessary in most hypothesis and values lessons. Either that, or persons occupying the teacher roles may feel impelled to talk a certain amount of time, perhaps because of the norms inherent in the situation.

Teacher questions were categorized into one of four possible categories: cognitive memory, convergent, divergent, and evaluative. Table 3 shows the proportions of total time in each category for each lesson type.

Table 3

PROPORTIONS OF TIME IN FOUR QUESTION CATEGORIES FOR THREE LESSONS

| Question Subcategory | Lesson Types Concept | Hypothesis | Values |
|---|---|---|---|
| Cognitive-memory | 6.03 | 4.37 | 3.57 |
| Convergent | 5.54 | 4.79 | 2.57 |
| Divergent | 0.95 | 3.92 | 1.72 |
| Evaluative | 0.57 | 0.57 | 6.52 |

The differences across lesson types form a reasonable pattern. Most cognitive-memory and convergent questioning appears in the concept lesson, with the proportions dropping in the hypothesis and values lessons, as we might expect. Divergent questions are most frequent in the hypothesis lesson, where the teaching task is to elicite several hypotheses to explain or answer the same phenomenon or question. Evaluative questions play a negligible part in the first two lesson types, but the higher proportion in the values lesson is indicative of the nature of that lesson task.

Student verbal behavior, as represented by means for each lesson type, follows the same general pattern as that of the teachers. Proportions of student talk in the response category are similar across the three lesson types: 18.51, 19.78, and 20.19 for the concepts, hypothesis, and values lessons, respectively. The student initiation category proportions are 8.71, 9.19, and 11.59, showing an anticipated rise in initiation as the lesson task becomes more open-ended.

Within the four subcategories of both student response and student initiation, the patterns resemble that of the teacher questions. Table 4 summarizes these proportions.

Table 4

PROPORTIONS OF TIME IN FOUR SUBCATEGORIES OF STUDENT RESPONSE AND INITIATION

| Student Response | Lesson Type Concept | Hypothesis | Values |
|---|---|---|---|
| Cognitive-Memory | 6.89 | 5.15 | 3.94 |
| Convergent | 8.94 | 8.14 | 4.34 |
| Divergent | 1.54 | 5.47 | 2.37 |
| Evaluative | 1.15 | 1.03 | 9.53 |
| **Student Initiation** | | | |
| Cognitive Memory | 3.88 | 4.11 | 3.18 |
| Convergent | 3.22 | 2.63 | 2.44 |
| Divergent | 0.92 | 1.97 | 1.72 |
| Evaluative | 0.69 | 0.48 | 4.24 |

The silence/confusion category is virtually unchanged across lesson types, with means of 11.92, 11.65, and 10.95.

Several interesting ratios can be derived from the basic interaction data. One, the proportion of teacher talk to all talk, has been presented above. Another ratio, having to do with teacher questions is the ratio of cognitive-memory and convergent to all questions. This is referred to as the 'closed question ratio. As expected, the ratio is a high .83 for the concept teaching lesson, and drops for the other two types, .67 for the hypothesis and .43 for the values lessons. The analogous student closed talk ratios are very similar in magnitude to the teacher closed question ratio: .84, .69, and .44, respectively.

Flanders computed a ratio of indirect to direct teacher behavior by dividing the sum of time spent in teacher categories 1-4 by all teacher talk. The higher the ratio, the more the teacher used "indirect influence", indicated by accepting and using student feelings and ideas, praise and questions, as compared with "direct influence," as represented by lecturing, directing, and criticizing. The values for this ratio are almost identical across lesson types: .40 for concept, .42 for hypothesis, and .40 for values. This result was not expected. It was predicted that the concept lesson would involve a more direct style than the other two types of lessons. Perhaps because all three lessons were taught in a discussion mode, the influence characteristic was actually controlled, in spite of the changes in task.

Two-stage behavior chains occur when the speaker shifts from one verbal category to another, or when teacher talk is followed by student talk, or vice versa. A teacher question followed by a student response is an example of the latter case, while teacher praise followed by teacher question is an example of the first type. The present data were searched for two-stage chains by selecting only those in a particular lesson type which comprised at least one per cent of the total lesson time. This move to limit the number under scrutiny seemed necessary because with the twenty-category system which was used in this research there are 300 possible two-stage chains.

The set of chains which begin with a teacher question is important to teaching in a discussion mode. In each of the lesson types there are several such chains containing one per cent or more of the total lesson time. The teacher question (cognitive memory)-student response (cognitive memory) chain was the most frequent such chain. Using the symbols from Figure 2, this is the 41-81 chain. Percentages in the three lesson types are 4.1 for concept, 3.0 for hypothesis, and 2.3 for values. Again, this pattern confirms that differences in teacher-student interaction vary as a function of teaching task. The 42-82 chain, which is teacher convergent question-student convergent response, has the same pattern: 3.3 per cent for concept, 2.8 for hypothesis, and 1.3 for values. The 43-83 chain, which is teacher divergent question-student divergent response, only appears in the hypothesis lesson with 1.3 per cent of time; in the other two types the chain accounts for only 0.5 and 0.7 per cent. Similarly, the 44-84 chain, teacher evaluative question-student evaluative response, accounts for 3.4 per cent of the values lessons, but only 0.3 per cent in each of the other two lesson types.

Interestingly, the teachers seemed to use lecture, as well as questions, as a soliciting tool. The 5-91 chain, teacher lectures-student initiates at cognitive memory level, appears in all three lesson types at just over the one per cent level. No other chain except one shows up as eliciting student initiations. This is the 02-9 chain, where 02 is unproductive silence/ confusion. In the concept lessons, the 02-92 chain appeared at the 1.0 per cent level (0.9 per cent for each of the others), while in the values lesson the 02-94 accounted for 1.4 per cent of time (0.2 per cent for the others). Students may be moving to "rescue" a teacher or another student from the embarrassment of confused silence.

Another frequently appearing chain in all three lesson types is the 91-5 or student initiated talk-teacher lecture. It occurred 1.1 per cent of the time in each of the lesson types. In fact lecture seems to be the typical response to student-initiated ideas in any of the four cognitive levels. No frequent chain starting with a student-initiated idea ends with any teacher behavior except lecture. In contrast, several chains starting with student <u>responses</u> end with either teacher praise or use of student ideas, which are thought to be "rewarding" verbal behavior. These teachers, therefore, tend to reward student response, while they tend not to reward student-initiated ideas. A similar pattern occurs when the <u>level</u> of student response or initiation is examined. In all three lesson types, the teachers gave more praise and used student verbalizations more if they were at the cognitive-memory or convergent levels, as compared with the divergent and evaluative levels.

Analysis of teaching task success. Multiple regression analysis was used to relate simultaneously clusters of predictor variables to the classrooms intermediate effects. Following a procedure outlined by Veldman, several regression models were compared with a random chance model and with one another to determine the relative contribution of variable clusters in explaining the outcome measures.[11] This general strategy is followed by Birkin in a recent volume.[12]

A complete report of the findings is not within the scope of this paper, but the most salient results can be discussed. For the concept teaching lessons, the first of two intermediate effects criteria is the number of examples of the concept or concepts being taught which are suggested by students. This variable was chosen because the ability to cite an example or instance of a concept is one way of judging from student behavior that they understand the concept to which the examples relate. The second intermediate effects criterion is the overall student response quality, especially focussing on student verbal cues which signal concept understanding.

The first two variable clusters which are analyzed with respect to these criteria are the marginal proportions in Flander's basic ten categories, and the marginal proportions in the expanded set of twenty categories. Using a number of examples as the criterion, the regression equations result in squared multiple R's (SMR'S) of .10 for the basic ten categories, and .17 when the additional ten categories are added. Neither the F-ratios of the individual models or of their difference indicated statistical significance at the .05 level. Results from the same predictor clusters while using student response quality are very similar (SMR's R's of .10 and .18.)

The third regression model includes the expanded twenty categories plus the nineteen cell proportions which exceeded .01. This increased the SMR from .17 to .33, or almost twice the variance accounted for. The F-ratio for this model is not significant at the .05 level.

The next variable cluster consists of the seven teacher trait variables, including five Conceptual Systems Test factor scores, the Social Science Skills Test score, and The Teacher Control score. Using number of examples as the criterion this combination results in a multiple R of only .20, and the SMC of .04. Interestingly, however, when overall student response quality is the criterion, the same seven variables combine for a multiple R of .47, and the SMC of .22, a much wider difference than encountered with the other variable clusters. The F-ratio for this model is significant at the .001 level. The reason for this striking anomaly is not apparent. The set of high inference rating variables, when entered into a regression equation, yield SMC's of .15 for the number of concepts criterion and .44 for student response quality, the latter significant at the .001 level. In both cases, clarity of concept precentation has the highest beta weight.

When all variable clusters described above are entered into a single multiple regression equation, the SMC is raised to .36 for the number of student examples criterion, and to .64 for the student response quality criterion, again with the latter significant at the .001 level. Close inspection of the latter equation reveals an interesting feature. The three highest betas are the following:

| Variable | beta |
|---|---|
| Category 5, lecture | .30 |
| Clarity of concept presentation | .51 |
| cell 5-5, steady state lecture | -.31 |

These results suggest that teacher explanation which is brief but clear is strongly associated with accomplishment of the intermediate classroom effects of concept teaching.

In summary, the cluster of high inference rating variables yielded the highest SMC (.44), with teacher traits second (.22), and the expanded twenty observation categories third (.18). By combining all variables in a single equation the SMC is raised to .64, with clarity of teacher explanation the strongest single trait associated with success in concept teaching.

In turning to the hypothesis and value analysis lesson types, it is helpful to summarize all the squared multiple correlations for the three lesson types and different criterion variables in one table. Table 5 provides this summary. It should be pointed out that the variables for the value analysis lessons do not include a frequency count similar to the number of examples or hypotheses variables as in the other two lessons. The student response quality variable is the only intermediate effects variable for the values lesson, therefore.

The pattern of SMC's for the hypothesis and values lesson are quite similar to those for the concept lessons discussed above. One of the two largest differences is in the uniformly small SMC's for the teacher trait variable cluster in the hypothesizing and values lesson, as compared with the .22 for the concept lessons. The second large difference lies in the substantially larger SMC's for the low-inference variable cluster in the hypothesizing and value analysis lessons.

If the three similar criterion variables -- the student response quality in each lesson -- are examined across the three lessons, the SMC's increase in all variable clusters, except for the teacher trait in the concept lesson as noted above. This systematic rise may be due to the increase of teaching experience over the semester's period. An alternative explanation lies in the complexity of the teaching tasks. The three types increase in complexity from the concept teaching to the values analysis teaching. It may be that success in teaching concepts may be accomplished through many different and perhaps unpredictable teacher behaviors, while the more complex tasks require a narrower range of behaviors for successful intermediate effects. In any case, the overall pattern of SMC's is quite regular, and the variables used in the analysis accounts for a surprisingly large proportion of variance in the intermediate success criteria. In general, the high inference rating variables are followed in relative predictive importance by low-inference observation variables, while the teacher trait variables come in a poor third.

The hypothesis lesson regression equation which includes the total variable set has some interesting features. Student divergent response has the highest beta of .58, which is quite reasonable since the criterion is number of student hypothesis, which presumably depend upon divergent student thought. Interestingly, however, the steady-state student divergent thought cell has a -.18 beta, suggesting that extended student verbalization can be disfunctional. The total praise

Table 5

SMC'S FOR VARIABLE CLUSTERS AND LESSON TYPES

| Variable Clusters | Concept | | Hypothesis | | Values |
|---|---|---|---|---|---|
| Criterion: | No. Examples | Student Quality | No. Hypothesis | St. Quality | St. Quality |
| Basic Ten Categories | .10 | .10 | .37*** | .28*** | .29*** |
| Twenty Expanded Categories | .17 | .18 | .50*** | .35*** | .38*** |
| Expanded Categories plus Transition-Cells | .33 | .31 | .56*** | .42*** | .49*** |
| Teacher Traits | .04 | .22*** | .02 | .02 | .09* |
| High Inference Ratings | .02 | .44*** | .27*** | .54*** | .64*** |
| Total Variable Set | .36 | .64*** | .62*** | .73*** | .82*** |

* = P < .05 These significance levels were derived from F-ratios computed
** = P < .01 as suggested by Veldman, opcit.
*** = P < .001

variable has a .28 beta, but the student convergent response - praise transition cell has a -.27 beta, suggesting that if teachers mistakenly praise convergent student thinking when a divergent task is called for, there is a corresponding decrease in hypothesis formation.

The values lessons total regression equation presents a challenging interpretation problem. Here are the variables with the largest beta weights:

| Variable | beta |
|---|---|
| Proportion of teacher evaluative question | -.55 |
| Proportion of student evaluative response | -.79 |
| Task clarity | .42 |
| T. evaluative question - S. evaluative response transition cell | .58 |
| Steady state S. evaluative response cell | .62 |
| Steady state T. evaluative question cell | .31 |

The pattern seems confusing and contradictory, in that total time in the evaluative categories receives a negative beta weight, while the corresponding steady state cells have positive betas. This may not be uninterpretable, however, Perhaps it is not what proportion of time the teacher or student spends in the evaluative realm of discourse, but rather the way in which the time is spent. Apparently, relatively small, concentrated verbal bursts in the evaluative realm, rather than extensive, prolonged verbal discourse is one key to success in Value analysis teaching tasks. Another key is the clairty with which the task is presented to students. These interesting findings need replication before basing practice upon them is possible.

Relation of low- to high-inference variables. The final question to be addressed has to do with how much overlap there is between low- and high-inference teacher behavior variables. Again, multiple regression techniques have been used to answer the question. Five rating variables, shown in Table 6, were entered into a regression equation with a set of 40 observation variables. The SMC's for each rating variable within each lesson type are also given in Table 6. In general, it can be

Table 6

SMC's of Low- to High-Inference Variable Regression Equation

| High Inference Teacher Variables | Concept | Hypothesis | Values |
|---|---|---|---|
| Task Clarity | .44** | .48*** | .45*** |
| Interesting Lesson introduction | .38* | .39** | .42*** |
| Enthusiasm | .38* | .24 | .35* |
| Business like behavior | .25 | .22 | .36* |
| Clarity of the problem or concept | .40* | .33* | .37* |

* = $P < .05$
** = $P < .01$
*** = $P < .001$

seen that by using the twenty category expanded Flanders observation system, from one-third to one-half of the variance in high inference variable can be accounted for in a regression equation.

Summary

Although the results reported above are tentative, they are encouraging for future studies which focus on both high- and low-inference variables as predictors of success in specific teaching tasks. Substantial proportions of criterion variance can be explained by both kinds of variables when such variables are clustered in multivariate analyses. Although the relative predictive power of high-inference variables appears higher than of low-inference variables, the latter also appear to overlap the former to a considerable extent. Finally, the teacher trait measurer such as cognitive complexity, social science intellectual skills, and attitudes toward teacher control are not useful in predicting success in intermediate classroom effects.

## NOTES

1. Borah Rosenshine and Norma Furst, "Research on Teacher Performance Criteria," in B. O. Smith, ed., Research in Teacher Education: A Symposium, Englewood Cliffs, New Jersey: Prentice-Hall, 1971, p. 42.

2. Ibid., pp. 56-8.

3. Bruce J. Biddle, "The Integration of Teacher Effectiveness Research," in Bruce J. Biddle and William J. Ellena, eds., Contemporary Research on Teacher Effectiveness, New York: Holt, Rinehart, and Winston, 1964, pp. 1-40.

4. A complete description of the basic ten categories can be found in Ned A. Flanders, Analyzing Classroom Behavior, Reading, Mass.: Addison, Wesley, 1971.

5. See the Aschner-Gallagher coding system described in Anita Simon and E. Gil Boyer, Mirrors for Behavior II, Volume A, Classroom Interaction Newsletter, Spring 1970, pp. 3-4 to 3-6.

6. O. J. Harvey, David E. Hunt, and Harold M. Schroder, Conceptual Systems and Personality Organization, New York: Wiley and Sons, 1961.

7. Clair M. Bowman, "Factors Influencing Pre-service Social Studies Teachers' Probing Behavior During Microteaching," paper delivered at the Annual Meeting of the National Council for the Social Studies, Boston, 1972.

8. Ibid.

9. Flanders, Analyzing Classroom Behavior, op cit, p. 101.

10. Arno A. Bellack, et al, The Language of the Classroom, New York: Teachers College Press, 1966, p. 55.

11. Donald J. Veldman, Fortran Programming for the Behavioral Sciences, New York: Holt, Rinehart and Winston, 1967, pp. 294-307.

12. T. A. Birkin, "Toward a Model of Instructional Processes," in Ian Westbury and Arno A. Bellack, eds., Research Into Classroom Processes: Recent Developments and Next Steps, New York: Teachers College Press, 1971, pp. 119-138.